UNCLASSIFIED

AD NUMBER
ADB378365
LIMITATION CHANGES
TO: **Approved for public release; distribution is unlimited.**
FROM: **Distribution authorized to U.S. Gov't. agencies only; Proprietary Information; 20 JUN 2011. Other requests shall be referred to Office of Naval Research, 875 N. Randolph St., Arlington, VA 22203-1995.**
AUTHORITY
NAVONR form 5510/7 dtd 8 Jun 2012

THIS PAGE IS UNCLASSIFIED

www.aptima.com
Boston ▪ DC ▪ Dayton

Security Classification: Unclassified

Final Report

Submission Date: March 08, 2012

Performance Measures for Improved Submarine Decision Making

(Aptima Job #1612)

Reporting Period: 06/20/2011 –03/08/2012

Prepared for:

Dr. William Krebs, Code: 342
Office of Naval Research
875 North Randolph Street
Arlington, VA 22203-1995
William.krebs@navy.mil

Prepared by:

Eric Jones, Jonathan Lansey, Fred Diedrich
Aptima, Inc.
12 Gill Street, Suite 1400
Woburn, MA 01801
ejones@aptima.com, jlansey@aptima.com, fdiedrich@aptima.com,

Distribution Statement A: Approved for public release: distribution unlimited.

Contract Number: N00014-11-M-0329
Contractor Name: Aptima, Inc.
Contractor Address: 12 Gill Street, Suite 1400, Woburn, MA 01801

This report was processed and delivered by Cynthia Devlin-Brooks, Contracts Specialist

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
03/08/2012

2. REPORT TYPE
Final

3. DATES COVERED *(From – To)*
06/20/2011 –03/08/2012

4. TITLE AND SUBTITLE
Performance Measures for Improved Submarine Decision Making

5a. CONTRACT NUMBER
N00014-11-M-0329

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

4. AUTHOR(S)
Eric Jones
Jonathan Lansey
Fred Diedrich

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Aptima, Inc.
12 Gill Street, Suite 1400
Woburn, MA 01801

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Dr. William Krebs, Code: 342
Office of Naval Research
875 North Randolph Street
Arlington, VA 22203-1995

10. SPONSOR/MONITOR'S ACRONYM(S)
ONR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution Statement A: Approved for public release: distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT *(Maximum 200 words)*
This effort focuses on the development, analysis, and evaluation of measurement tools for submarine navigation teams, including support of Future Naval Capabilities efforts, and in particular, the 10-02, Adaptive Training for Submarine Navigation and Piloting (AT-SNAP) program. The primary challenge was to explore methods to automatically assess team processes in ways that are not fully dependent on instructors or observers, as is presently the case. This effort aimed to identify techniques that can automatically assess various aspects of the performance of submarine navigation teams. The Aptima team, while supporting Dr. David Kern (AT-SNAP program manager) and Sandia National Laboratories, studied the use of Sociometric Badges (produced by Sociometric Solutions, Inc.) to supplement data collection efforts. The Sociometric Badges are unobtrusive pieces of hardware that employ multiple sensors to collect various types of data as teams of people interact in complex mission environments. The badges were used to collect data as submarine crews performed exercises in Submarine Piloting and Navigation trainers during a two-day study at the Naval Submarine School in Groton, Connecticut. Although this effort is exploratory, preliminary results suggest a number of findings that speak to the benefit of Sociometric Badge technology when applied to the undersea warfare domain.

15. SUBJECT TERMS
Submarine Team Performance, Sociometric Badge, Automated Evaluation, Piloting and Navigation, Team Processes

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Eric Jones
Unclassified	Unclassified	Unclassified	UU	47	19B. TELEPHONE NUMBER (include area code) 781.496.2426

Computer Generated

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std Z39-18
Microsoft Word

TABLE OF CONTENTS

List of Figures

Figure **Page**

List of Tables

Table **Page**

1. Summary

The work reported here focuses on the development, analysis, and evaluation of measurement tools for submarine navigation teams , including support of Future Naval Capabilities efforts, and in particular, the 10-02, Adaptive Training for Submarine Navigation and Piloting (AT-SNAP) program. As part of this program, Aptima provided support to the Office of Naval Research (ONR) in collaboration with Dr. David Kern (program manager) and Sandia National Laboratories (SNL). The primary challenge of this effort was to explore methods to automatically assess team processes in ways that are not fully dependent on instructors or observers, as is presently the case, for instance using the Continuing Training Support System (CTSS). For AT-SNAP, for example, these measures will ultimately be used to automatically assess team state in the context of piloting and navigation, thereby enabling performance feedback and adaptation of training in order to promote learner-centered instruction. Given the focus on team processes, we were particularly interested in assessing aspects of communication and information transfer (e.g., see Smith-Jentsch et al., 1998). The challenge addressed here is to identify techniques to automatically assess these aspects of performance, focusing on submarine navigation teams in context of surface transit.

More specifically, Aptima studied the use of *Sociometric Badges* to supplement data collection efforts for the AT-SNAP program, and worked with Dr. Kern and SNL to begin exploring the applicability of these devices to the piloting and navigation domain. The Sociometric Badges, produced by Sociometric Solutions Inc. (SSI), are small, unobtrusive pieces of hardware that are worn around a person's neck and employ multiple sensors to collect various types of data as teams of people interact in complex mission environments. The types of data that are recorded include artifacts of speech, face-to-face interactions, and the proximity of people with respect to one another. Gross body movements are also recorded (for example, whether or not a person is walking or running), though these data were not explored in this effort. The badges were used to collect data as submarine crews performed exercises in Submarine Piloting and Navigation (SPAN) trainers during a two-day study at the Naval Submarine School (NSS) in Groton, Connecticut. The data were then analyzed to assess the ability of the Sociometric Badges to automatically and reliability detect behaviors that correlate to team performance.

Although this effort is exploratory, preliminary results suggest a number of findings that speak to the benefit of Sociometric Badge technology when applied to the undersea warfare domain. The Sociometric Badges seem to be uniquely suited to assessing the state of submarine teams. For example, volume as captured by the Sociometric Badges is a promising way to detect what the team is doing (e.g., where their focus of attention is), and determine what they should be doing (e.g., patterns in volume that correspond to better execution of cyclic routines; tension that should exist given certain mission conditions). The data that are collected from the infrared (IR) sensors can be used to map control room activity by capturing the frequency of crewmember interactions. This data can also be used to show how the crew tends to move around the control room during the mission. Some challenges to this technology remain, but as these capabilities mature, there will be additional opportunities to advance this work. Overall, unobtrusive measurement of team processes using the Sociometric Badges provides a novel and promising step in automated submarine team assessment.

2. Introduction

The work reported here focuses on the development, analysis, and evaluation of measurement tools for submarine navigation teams , including support of Future Naval Capabilities efforts, and in particular, the 10-02, Adaptive Training for Submarine Navigation and Piloting (AT-SNAP) program. As part of this program, Aptima provided support to the Office of Naval Research (ONR) in collaboration with Dr. David Kern (program manager) and Sandia National Laboratories (SNL). The primary challenge of this effort was to explore methods to automatically assess team processes in ways that are not fully dependent on instructors or observers, as is presently the case, using for instance the Continuing Training Support System (CTSS). For AT-SNAP, for example, these measures will ultimately be used to automatically assess team state in the context of piloting and navigation, thereby enabling performance feedback and adaptation of training in order to promote learner-centered instruction. Given the focus on team processes, we are particularly interested in assessing aspects of communication and information transfer (e.g., Smith-Jentsch et al., 1998). The challenge addressed here is to identify techniques to automatically assess these aspects of performance, focusing on submarine navigation teams in context of surface transit.

More specifically, Aptima studied the use of *Sociometric Badges* to supplement data collection efforts for the AT-SNAP program, and worked with Dr. Kern and SNL to begin exploring the applicability of these devices to the piloting and navigation domain. The Sociometric Badges, produced by Sociometric Solutions Inc. (SSI), are small, unobtrusive pieces of hardware that are worn around a person's neck and employ multiple sensors to collect various types of data as teams of people interact in complex mission environments. The badges were used to collect data as submarine crews performed exercises in Submarine Piloting and Navigation (SPAN) trainers at Naval Submarine School (NSS) in Groton, Connecticut. The data were then analyzed to assess the ability of the Sociometric Badges to unobtrusively, automatically, and reliability detect communication and coordination behaviors that correlate to team performance.

The analyses and findings reported here are intended to be exploratory. The intent of the study, which involved a small sample of two teams conducting training scenarios that were not influenced by the research team, was to generate initial data that could be used to explore the potential applicability of the data collection methodology to the submarine domain. Accordingly, while we present sample exploratory findings, these findings are not conclusive and are intended to guide further study, refinement, and validation as the AT-SNAP program continues, should the program seek to employ the Sociometric Badges to assess team coordination.

3. Methods, Assumptions, and Procedures

3.1 Sociometric Badge Technology

The Sociometric Badges, produced by Sociometric Solution Inc. (SSI), are small, unobtrusive pieces of hardware that are worn around a person's neck (Figure 1). They are intended to be worn by multiple people during missions or exercises, and they employ a variety of onboard sensors to collect data as teams of people interact. Each badge contains microphones, infrared (IR) detectors, accelerometers, and Bluetooth transceivers which are all connected to a computing system. There are two microphones, one on the top side of the badge to sense the voice of the person who is wearing it, and one on the front side to pick up sound from people to whom the wearer is speaking (see Figure 1, left). Raw audio is not recorded, but rather the signal is compressed in real-time into a rolling average of amplitude. This recording technique not only avoids privacy and security issues, but is also essential to conserving power (as of now, the badges can run continuously for over 40 hours on a full charge). There is a limited aperture IR transceiver, which can sense when it is aligned with the IR transceiver of another badge. This essentially records when two people are facing each other, or in other words, interacting in some way based on the context of activity being observed. To capture over-the-shoulder interactions, as are seen during navigation and piloting, badges are placed on different workstations to detect both when a crewmember is sitting in a particular seat and when someone walks up from behind. The Bluetooth sensor, which sends out a signal and receives a reply from all badges within range, has a measure of signal strength that, in theory, can be used to estimate distance between them. The accelerometers detect motion, and are currently used to sense gross body language (e.g., running vs. walking), but were not yet explored in this effort.

Figure 1: A diagram of a Sociometric Badge (left), a person wearing a Sociometric Badge (center), downloading data from the Sociometric Badges (right). Pictures are supplied by the SSI Sociometric Badge User Manual (Sociometric Solutions, Incorporated, 2011).

The badges are designed to be simple to use in that after they are put around the neck and turned on (Figure 1, center), no further interaction on part of the wearer is necessary. The downloading of recorded data is performed easily and automatically when the badges are plugged into a computer that is running the SSI software (Figure 1, right). Furthermore, several badges can be downloaded at the same time, greatly reducing the amount of time needed to perform this step. Note that several hours of data take no more than 30 minutes to fully upload.

In addition to these core functions, in other work, Aptima and SSI are developing several higher-level features that build on the existing sensing capabilities of the Sociometric Badges. For example, consider that the audio signal activity level is currently recorded as a rolling average of amplitude. This can be used to detect changes in volume, such as talking at a normal level versus shouting. With future versions of the SSI software, researchers will be able to calculate the percentage of speaking, listening, silence, and overlap in conversation throughout an exercise for a particular participant. From this, an overall *dominance score* can be calculated relative to the other participants for each session along with the average speaking segment length and the average pause lengths. One additional calculation that is of particular interest is the *turn-taking adjacency matrix*, which includes the total number of conversational "turns" taken between each individual as well as an overall influence score for each session. As these capabilities continue to be developed, it is expected that these features may be highly informative about the nature of interactions between submarine crews who are required to perform specific litanies that emphasize patterns of turn-taking, confident tones, and rhythms.

3.2 Pilot Study

In December 2011, employees of Aptima traveled to Sandia National Laboratories (SNL) to participate in a working meeting. The goals of this trip were to familiarize the team with the Sociometric Badge technology, to explore how this technology could be incorporated in a data collection event, and to collect example data in a pilot study to begin exploring the capabilities and limitations of the badges. During this meeting, several tests were run to verify the conditions under which the badges performed well, and various conditions were simulated that were expected to be encountered when data were collected with a real crew. The range of distances and angles over which the IR detectors worked were examined in order to better understand the range of face-to-face interactions that the badges will be able to detect. It was determined that a badge could receive signals from other badges that were 3 feet away and off-axis by an angle of up to 55 degrees (0 degrees being the two badges directly facing each other). At distances greater than 3 feet, this angle decreased gradually. When perfectly aligned (0 degrees) two badges needed to be within approximately 5-6 feet to reliably detect each other. These ranges were considered sufficient to capture a face-to-face interaction between two crewmembers. Table 1 below shows an example of data collected from the IR detectors of badge number 376 ("Badge Receiver" column). It detected that over a period of several seconds it was facing badge number 434 ("Badge Sender" column). For the infrared sensors, one second roughly maps to one row: this data recording rate can be adjusted to be faster, though at a cost of battery life.

Table 1: Sample data showing interactions between two Sociometric Badges based on IR sensor detection.

Badge Reciever	Timestamp	Badge Sender
376	12/15/11 16:20:26.765	434
376	12/15/11 16:20:27.726	434
376	12/15/11 16:20:28.757	434
376	12/15/11 16:20:29.738	434
376	12/15/11 16:20:30.779	434
376	12/15/11 16:21:00.510	434

The data that are recorded from the Bluetooth transceivers is similar to the IR data (badges that return the signal are recorded in a similar format); however, there is an additional parameter of

Received Signal Strength Indication (RSSI; see Table 2). In trial experiments, the manner in which this signal strength varies with distance and the presence of occluding obstacles was investigated. The badges were first scattered around an area to simulate crewmembers positioned at different distances and among different objects (such as walls and furniture). Then, similar tests were run with the badges spaced closely together while hanging on the backs of office chairs to simulate the manner in which they will be worn by a person. It may be possible to reconstruct each person's approximate location from this data, though early attempts at looking at data from a single badge were not conclusive given the range of distances that are expected within submarine training environments. In the context studied, the average RSSI between two badges tended to be higher when they were within 1 foot of each other, but the signal quickly dropped off as that distance increased. In addition, the variability in the RSSI value was large when observed over a period of 30-60 minutes, implying that at any point in time, it may be unclear where the other badges may be positioned. It is important to note that these were early attempts to explore this application of the data, and with further exploration and tuning of various settings and features, it is likely that this accuracy will increase. At that time, RSSI may be able to be fully exploited within the undersea domain.

Table 2: Sample Bluetooth data showing which badges were detected and the Received Signal Strength Indicator (strength increases as the numbers become more positive).

Badge Reciever	Timestamp	Badge Sender	Rssi
376	12/15/11 16:18:48.632	434	-78
376	12/15/11 16:18:49.043	445	-63
376	12/15/11 16:18:49.203	444	-85
376	12/15/11 16:18:50.675	446	-71
376	12/15/11 16:19:41.148	446	-72
376	12/15/11 16:19:41.828	445	-60

In addition to these basic tests, the SubSkillsNet simulation environment was used to perform a round of contacts (ROC) exercise while wearing the badges. Badges were also secured to the different workstations, as they would be in the SPAN trainer. During each ROC, the project team worked together to use the simulated periscope and radar to scan the immediate environment and report bearings and ranges of various contacts to the "instructor." Below is an example of audio data for badge number 376. It includes the amplitude of the signal, the standard deviation and the minimum and maximum volume (Table 3). It was also determined that when a person is looking over the shoulder of an operator for 30-60 seconds, the badge is likely to pick up his presence using the IR data. If the time is shorter, however, then the likelihood of detecting this interaction will decrease. Interactions that are much less than 30 seconds may or may not be detected during an actual exercise. Regardless, the team was confident that the Sociometric Badges would effectively capture most of the essential expected crew behaviors.

Table 3: Sample audio data that was collected during a round of contacts (ROC) exercise.

Badge ID	Timestamp	Amplitude	Stddev	Min	Max	Mean
376	12/15/11 16:18:33.714	0.000124	0.002212	-0.04224	-0.02362	-0.03296
376	12/15/11 16:18:34.248	6.62E-05	0.001328	-0.03928	-0.02686	-0.0331
376	12/15/11 16:18:34.984	3.23E-05	0.001184	-0.03729	-0.02866	-0.03294
376	12/15/11 16:18:35.496	7.18E-05	0.001766	-0.04214	-0.02478	-0.03307
376	12/15/11 16:18:36.009	5.43E-05	0.001261	-0.03699	-0.02948	-0.03297
376	12/15/11 16:18:36.522	5.73E-05	0.001221	-0.03708	-0.02902	-0.03309

Overall, these preliminary data suggested that the Sociometric Badges would provide a number of novel capabilities that could supplement future data collection efforts. The conditions under which the badges successfully collected various types of data (e.g., the maximum distance at which IR pings were recorded, the sensitivity of the microphones) were determined to be relatively well-matched to the conditions that were anticipated during the training exercises. Some limitations of the technology were identified, but again, with further exploration and simple modifications to the settings, it is believe that these limitations could be addressed.

3.3 Naval Submarine School (NSS) Data Collection Overview

Following this intial pilot testing, the Sociometric Badges were used during a two-day data collection event that took place in early in 2012. During these two days, two different submarine crews were observed as they performed training exercises in two different Submarine Piloting and Navigation (SPAN) trainers at Naval Submarine School (NSS) in Groton, Connecticut. One crew was more experienced, and the other less experienced, as determined by the opinion of the subject matter experts who were observing the exercises. The more experienced crew, observed on Day 1, consisted of a Navigator (Nav), Assistant Navigator (ANAV), Quarter Master of the Watch (QMOW), Fathometer Operator, Radar Operator, Secondary Voyage Management System (VMS) Operator, Bearing Recorder, and Deck Log Recorder. In addition, various stations were instrumented with badges, including the Periscope, Fathometer, Primary VMS, Secondary VMS, and Radar Stations. The less experienced crew was observed on Day 2 and included similar personnel and equipment, except that the Nav and ANAV roles were performed by the same person, there was no Deck Log Recorder, and two crewmembers operated the Periscope (Periscope Operators A and B). Day 1 was divided into 2 sessions to focus on two different training scenarios: “Session 1” was considered an “easy” scenario, and “Session 2” was harder (according to the instructors and subject matter experts). Day 2 consisted of a single session, “Session 3,” which focused on a single training scenario the entire time.

In addition to the data that were collected using the Sociometric Badges, video of the exercise was recorded, though without audio. The purpose of the video was to provide a reference with which to corroborate the badge data, if necessary, and is not reported here. Because several of the computer screens in the SPAN trainer displayed classified information, this video was classified and treated accordingly. The project team also collected a series of notes in an effort to gauge how well each crew performed. Cross-track error (abbreviated “XTE”) is the difference between the actual position of the ship and the desired position as set in the VMS. This value changed continuously, but was recorded at regular intervals of time to provide an indication of the team’s ability to navigate a scenario. On both days, the crews engaged in different types of cyclic routines during which they practiced formal litanies at regular intervals to communicate ownship status and sensor information. These routines are integral to piloting and navigation tasks, and the manner in which they are conducted is considered an indicator of the skill and experience of the crew. Finally, a running log of scenario events was recorded so that interesting behavior in the data could be cross-referenced with mission context (see *Appendix A: Experimentation Forms* for the various templates that were used during the study). All of these notes were intended to be cross-checked with the Sociometric Badge data, to see if measured operator state corresponded to specific types of events, scenario difficulty, and/or the performance of the crew.

Before the start of each exercise, the project team delivered a short in-brief to the crew that explained the goals of the study and an overview of what we were asking them to do. Each crewmember who was part of the piloting and navigation team was then provided an informed consent document to review and sign. When the forms were signed, the crewmembers were given a Sociometric Badge to wear for the duration of the exercise. Each badge number was listed with its corresponding position on a reference sheet that was used when analyzing the data. Each badge was turned on prior to the study, and remained on until the exercise was complete. When the crew was finished training, each badge was retrieved and each crew member was asked to complete a survey to assess how well they felt they performed individually and as a team (see *Appendix B: Self-Reported Performance Survey*). The badge data were then downloaded to a computer where they could be analyzed.

4. NSS Study Results and Discussion

The following section describes an analysis of the various types of data that were collected using the Sociometric Badges, including: the volume level as recorded by the microphone, the line-of-sight interactions as recorded by the IR sensors, and the presence of surrounding badges as recorded by the Bluetooth capabilities. Note that these data are not conclusive, but exploratory given the context in which they were collected.

4.1 Volume Data

Subject matter experts in the Submarine domain have indicated that the volume and amount of discussion within the control room is an indicator of team performance, and that this volume should not exceptionally high or low for a high-functioning team under typical circumstances (Chester et al., 2011; Jones et al., 2010). To explore this notion further, data from the Sociometric Badges were examined to see how the recorded volume changed over time and with respect to different scenario events. Cross-track error was recorded in roughly 5-minute intervals during Day 1, but not during Day 2 because the nature of the training was different (training on this day emphasized lower-level tasks such as proper execution of litanies, and therefore did not stress this metric). On Day 1, when cross-track error was close to “0,” the team was performing well, and when it diverged greatly left or right, the team was performing poorly. Different scenario events challenged the team in different ways, and included, for example, changes in course, the loss and gain of different sensors, and various other piloting and navigation procedures. Changes in volume were examined for individual crewmembers and also for the crew as a whole. As volume changed during the course of the scenario, there were several opportunities to see how these data could be indicative of the team’s current state. All volume is captured by a digital microphone, and therefore measured in units of Decibels Relative to Full Scale (DBFS). DBFS is calculated using both the dynamic range of the microphone and the digital signal that is output as the microphone picks up sound.

4.1.1 Individual and Team Volume

Figure 2 illustrates the cross-track error (the graph on the left) with respect to the total volume recorded by the badges on all crewmembers for Day 1, Session 1 (the bar on the right). The cross-track error values (measured in feet) to the left of the dotted line are left of the desired course, and those to the right are right of the desired course. Lighter colors (bright yellow) in the volume bar correspond to louder total volume. A few minutes prior to the largest cross-track error value, the crew can be seen reaching what appears to be the loudest volume level of the session. This would suggest that as the ship was about to deviate greatly from the desired course, the volume level of the discussion increased. Similarly, both instances in which the cross-track error returns to 0 are preceded by periods of relative quite (as indicated by the dark red bars). One possible explanation is that the crew had taken actions to right the ship, and confident of their actions, allowed those changes to take effect with minimal discussion.

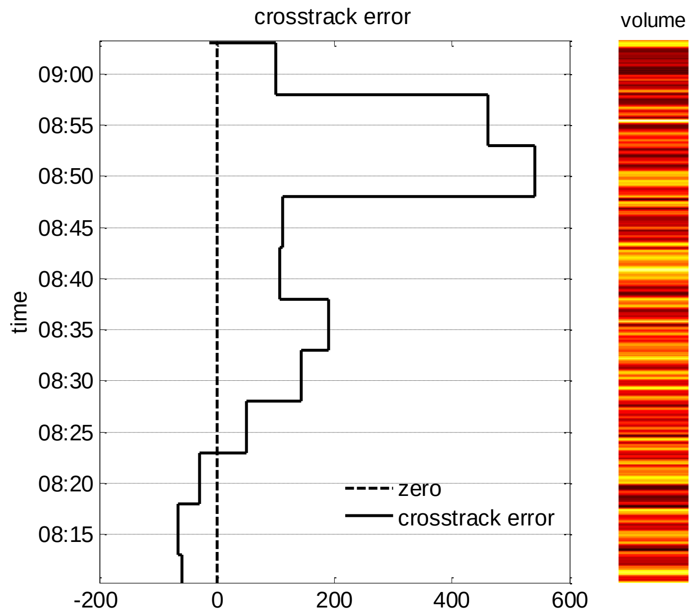

Figure 2: Total recorded volume (bar on the right) with respect to cross-track error (measured in feet) for Day 1, Session 1. Lighter colors (bright yellow) correspond to louder total volume (measured in DBFS). Time is measured here in minutes.

Figure 3 shows a similar graph for Day 1, Session 2, in which similar findings are not observed. There are no distinct periods of calm that appear to match with reductions in cross-track error, and likewise no periods of distinct loudness when this error was high. However, this exercise was different than the one experienced by the crew in the first session, which could explain the differences between the data. While the first session was performed in the open ocean on approach to a port, the second session not only included this approach, but also entry into the port and tight maneuvering in a narrow channel. This more difficult situation was much less tolerant of cross-track error, and therefore the level of stress in the control room was observed to be fairly high. This may be consistent with the relatively loud level of noise that was measured throughout this exercise.

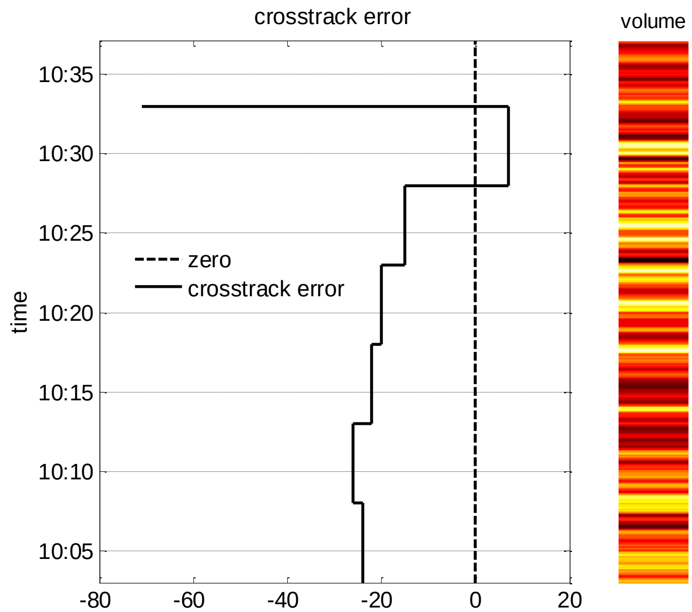

Figure 3: Total recorded volume (bar on the right) with respect to cross-track error (measured in feet) for Day 1, Session 2. Lighter colors (bright yellow) correspond to louder total volume (measured in DBFS). Time is measured here in minutes.

Recorded volume can also be plotted for individual crewmembers, as seen in the selected examples in Figures 4-8. In the following figures, the recorded volume of the Bearing Recorder, Navigator, Radar Operator, and Fathometer Operator were plotted over time and annotated with concurrent scenario events. The time (x-axis) does not always display consecutive minutes because of rounding errors in the calculations that are used to display them on the plots. Although they are approximate, they are accurate within 30 seconds, which is a shorter time scale than that by which the scenario events transpired. It is also important to note that the volume is recorded as a rolling average taken over a 32-second time window. This means that every point plotted is the average volume of the 16 seconds before and after it. This technique tends to minimize high frequency noise, and effectively display the underlying signal.

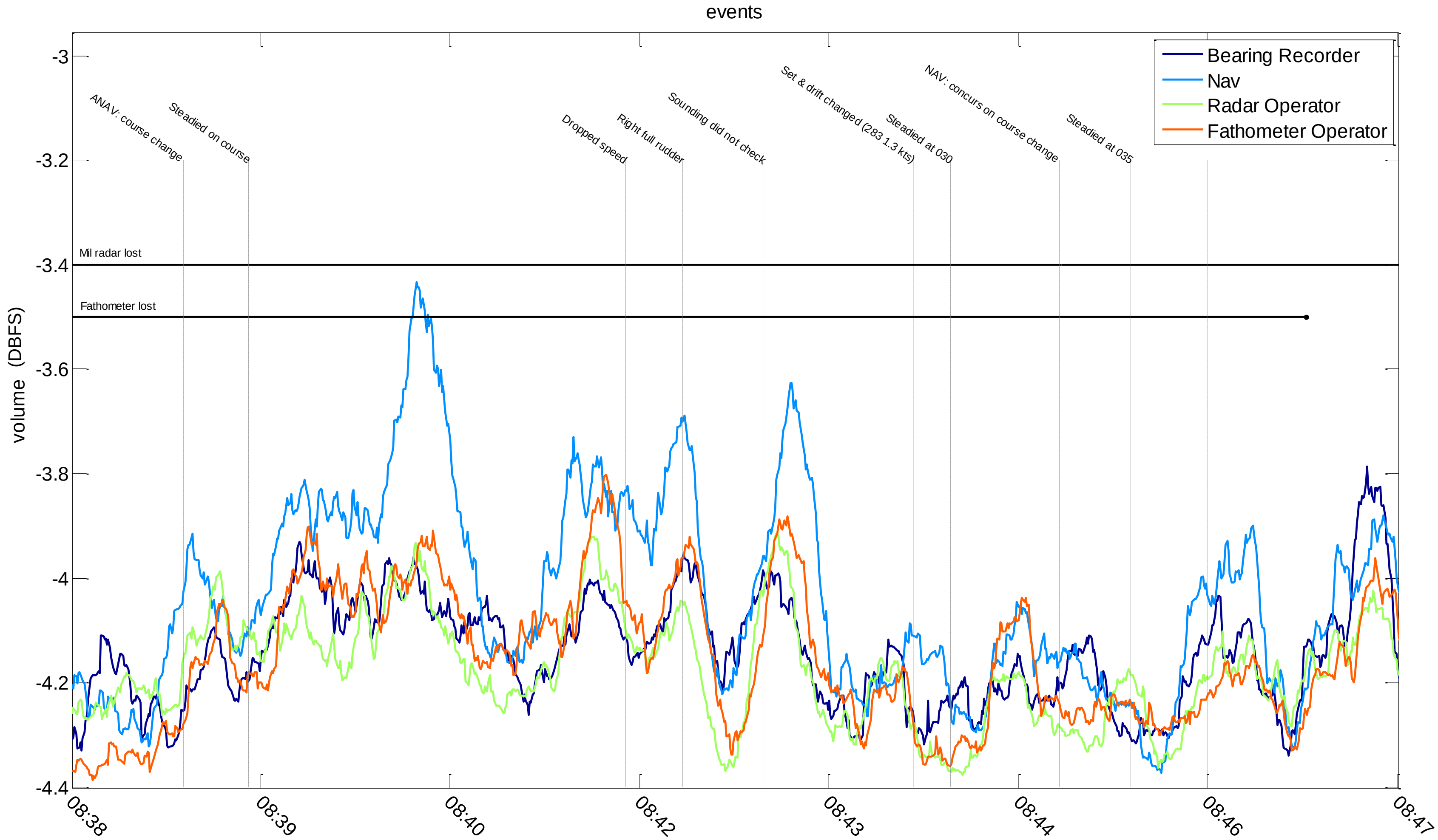

Figure 4: Individual recorded volume of the Bearing Coordinator, Navigator, Radar Operator, and Fathometer Operator for a selected example during Day1, Session 1. This example illustrates a response to a tense situation.

In Figure 4, the crew from Day 1 (Session 1) is faced with a tense situation that includes the loss of two sensors (Fathometer and Military Radar) and increasing cross-track error. The Nav can be seen loudly giving commands that precede actions by the crew to reduce speed and correct the ship's course. Similarly, when the sounding did not check, the volume in the control room appeared to increase momentarily, followed by a decrease to a lower, steadier state.

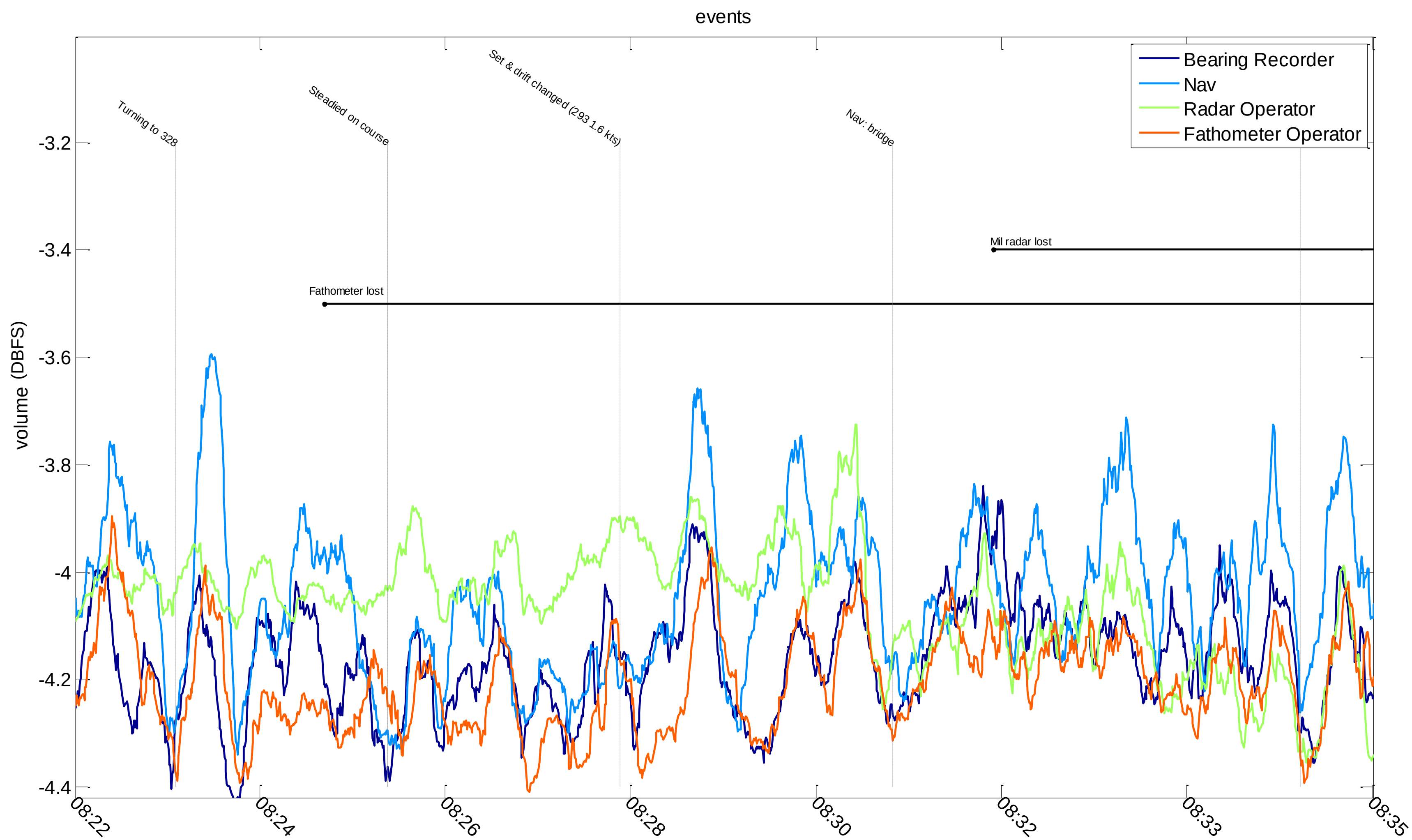

Figure 5: Individual recorded volume of the Bearing Coordinator, Navigator, Radar Operator, and Fathometer Operator for a selected example during Day1, Session 1. This example illustrates the volume when the ship is steadied on course.

In Figure 5 (from Day 1, Session 1), the crew is facing some challenges but is mostly steady and on course. This example represents a period of relative calm, without the same distinct peaks of volume seen in more tense situations. There is still a degree of punctuated discussion that, with more opportunities to observe the crew could potentially establish a baseline level.

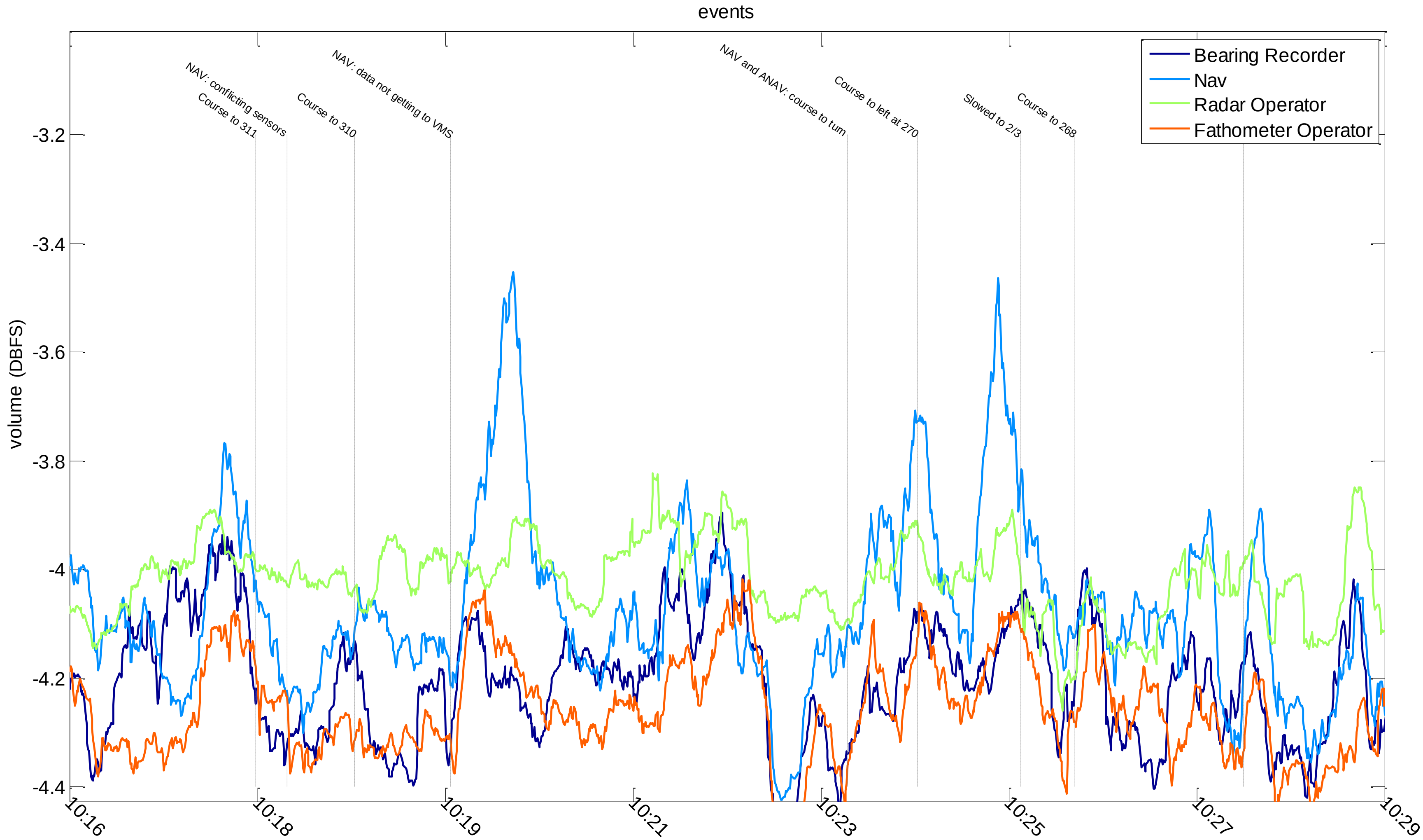

Figure 6: Individual recorded volume of the Bearing Coordinator, Navigator, Radar Operator, and Fathometer Operator for a selected example during Day1, Session 2. This example illustrates a tense situation in which the Nav instructed the crew.

The example in Figure 6 is from Day 1, Session 2. Here, the scenario was generally more difficult than the Session 1 exercise because of the crew was required to maneuver precisely through the channel. In this instance the Nav can be seen loudly instructing the crew (seen by the two high peaks at ~10:19 and ~10:25).

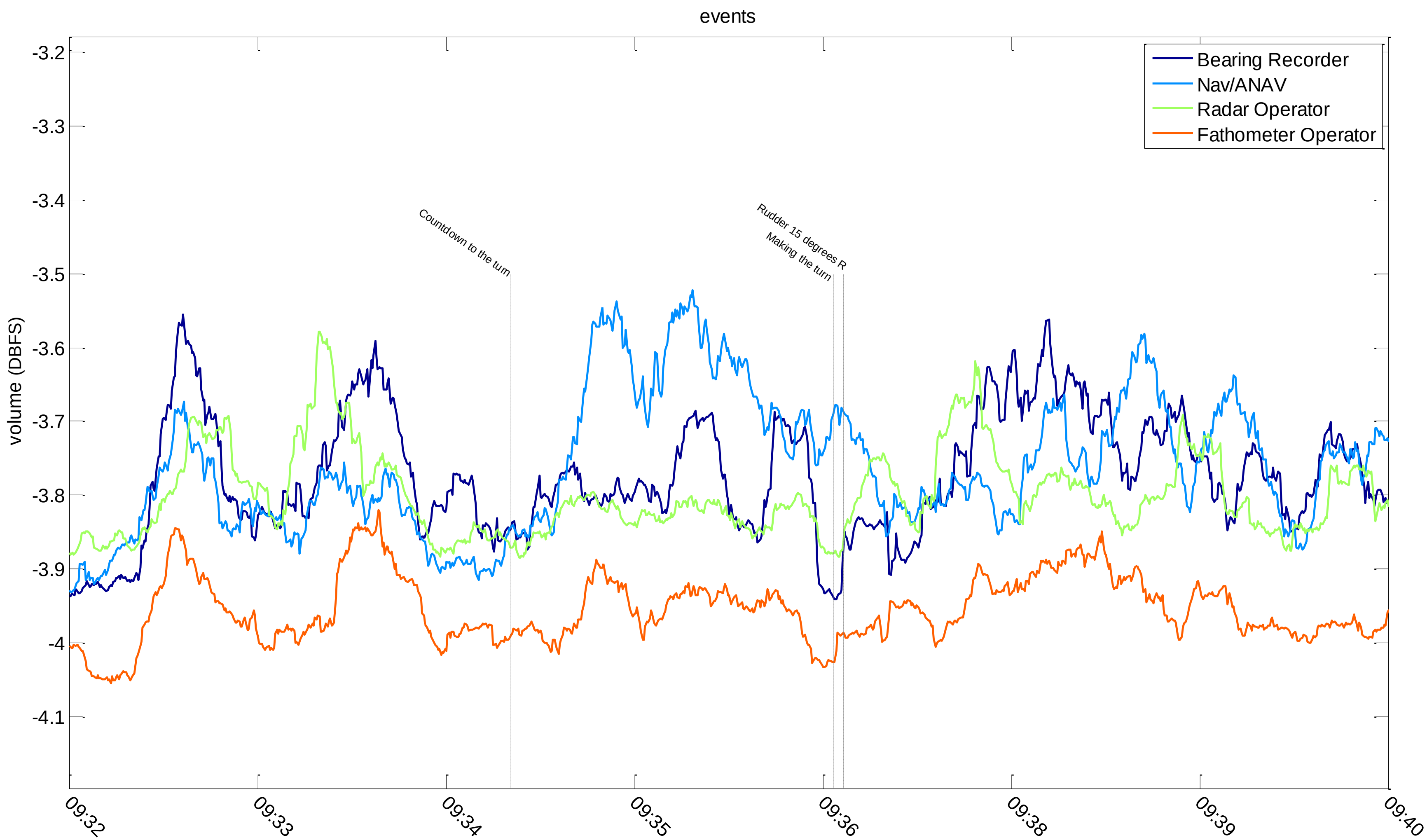

Figure 7: Individual recorded volume of the Bearing Coordinator, Navigator/Assistant Navigator, Radar Operator, and Fathometer Operator for a selected example during Day2, Session 3. This example illustrates a routine process (turning).

Figure 7 displays an example from the crew on Day 2, Session 3. Training on this day focused on more basic skills such as the execution of various cyclic routines. This practice can be seen in the patterns increasing and decreasing volume that seem to be passed from one crewmember to another. As the ship prepares to make a turn, the Nav can also be seen giving instructions in preparation.

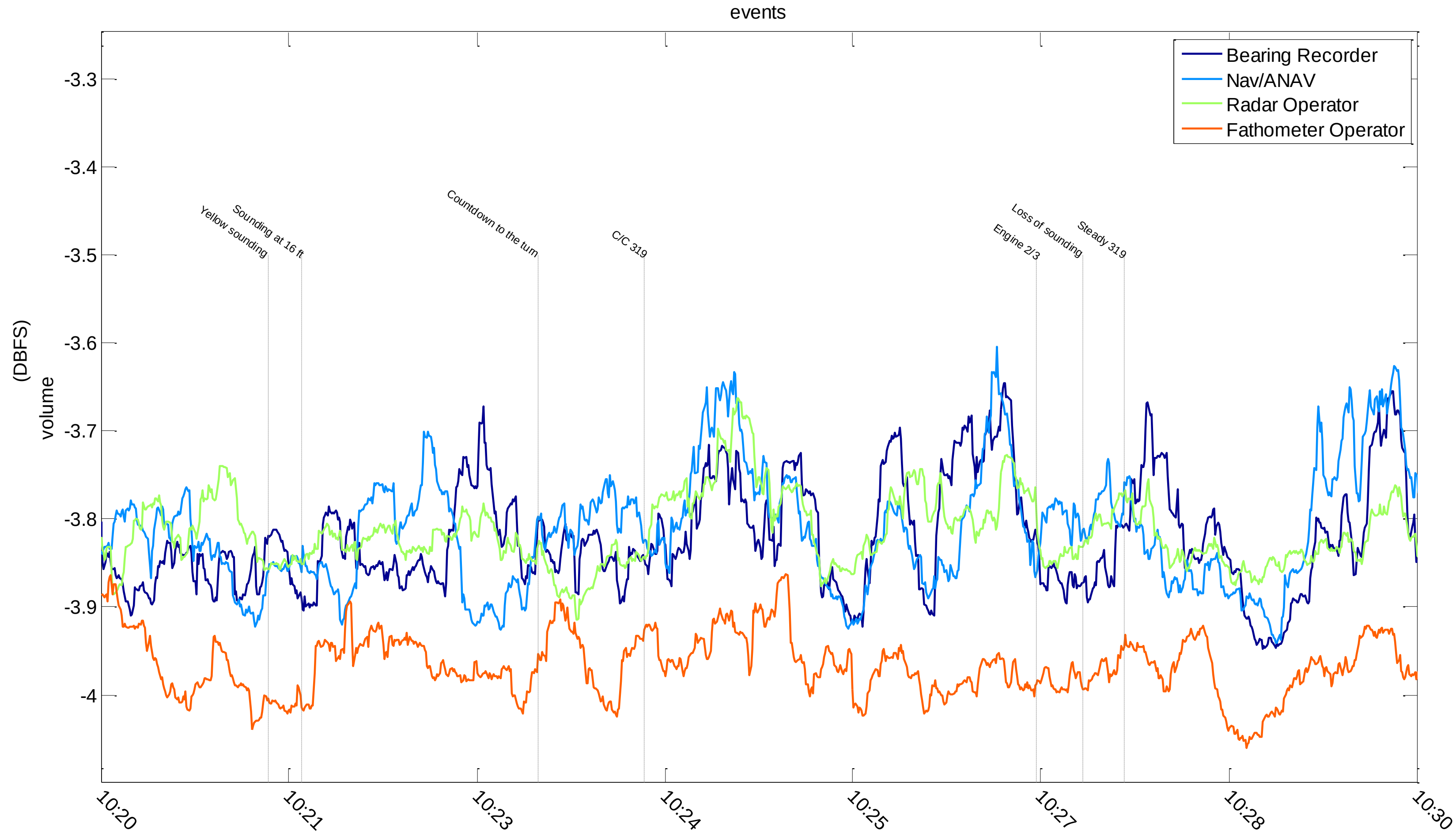

Figure 8: Individual recorded volume of the Bearing Coordinator, Navigator/Assistant Navigator, Radar Operator, and Fathometer Operator for a selected example during Day2, Session 3. This example illustrates the crew's reaction during a tense situation.

The last selected example (Figure 8) is also from Day 2, Session 3. Here, the crew found themselves in a tenuous situation where they received a yellow sounding, and then a sounding at 16 feet shortly thereafter. Although this was a serious situation that required immediate attention, the lack of change in volume (i.e., the recorded volume was consistent) may have that indicated either that the crew was not cognizant of the situation or was focused was focused in other areas, such as the process of litanies rather than position of ship per se given their training objectives.

Overall, Figures 4 to 8 suggest that volume escalates and modulates when commands and high level guidance are given, which could be used to infer crew state and patterns of dialog over time. For example, more direct, forceful, and loud commands could be indicative of effective interaction, or the presence or absence of commands at particular times could be considered indicators of performance as well. Similarly, the presence or absence of tension, as measured by volume may be a potential indicator of team state as well. Volume seems to map to the amount of tension in the control room, in that when the scenario was difficult, volume tended to increase. This increase in volume was seen more clearly with the more experienced team at difficult moments in the scenario. On the other hand, when similar situations were faced by the less experienced team, volume did not change dramatically, perhaps reflecting their training emphasis/focus. Collectively, these data suggest that patterns of changes in volume may be useful for discerning team state, although further analyses are needed as the AT-SNAP program evolves.

4.1.2 Volume during Cyclic Routines

Each team on Days 1 and 2 participated in cyclic routines (e.g., visual, radar, GPS only, etc.) that were intended to practice the various procedures as they would be performed when underway. More experienced, skilled crews are recognized as performing these cyclic routines over a short amount of time with crisp litanies and confident tones. Less experienced teams may take longer to complete each exercise, and the flow of information among the navigation and piloting party may not be as smooth as it could be. By looking at the volume data that is recorded from the Sociometric Badges, it is possible to see differences in the sound that is recorded during cyclic routines when performed by less or more experienced teams. Referring to Figures 9 and 10, all of the cyclic routines for Days 1 and 2 were normalized to a common starting point, as seen by the vertical dashed line (the time at which each routine started and ended was recorded throughout the observation; the start time is used to define this common starting point). Then, from this point of reference in time, the average volume that was recorded for each team was calculated (the solid black line), as well as the standard deviation around that average (the gray waveform). The resulting graph is a high-level overview of volume across all the cyclic routines.

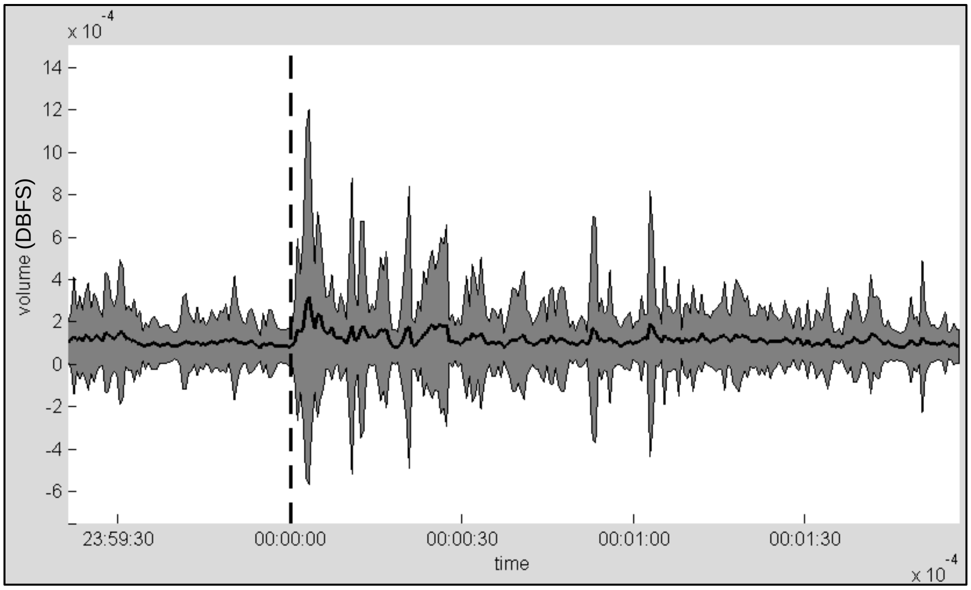

Figure 9: Average total volume (black line) and standard deviation (gray waveform) for all cyclic routines for Day 1 (Sessions 1 and 2).

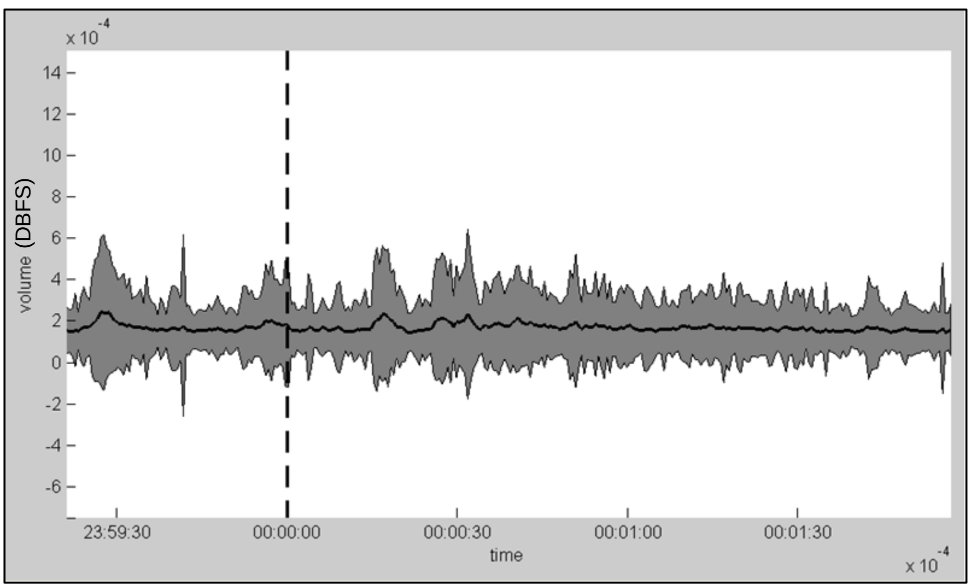

Figure 10: Average total volume (black line) and standard deviation (gray waveform) for all cyclic routines for Day 2 (Session 3).

There are several differences that are immediately apparent between the two teams. In Figure 9, the more experienced team exhibits a number of well-formed peaks in volume, which may be representative of the skill with which they were able to consistently follow the litany and take turns. By comparison, the less experienced team in Figure 10 displays a waveform that is much less defined. The more experienced team also had more variation in the recorded volume, meaning that there were likely times that the team was sharply performing the litany loudly and confidently. The less experienced team did not experience much variation, and did not reach similarly high levels of volume. One caveat is that the number of cyclic rounds performed on Day 1 (approximately 20) differed from the number performed on Day 2 (approximately 50). Also, because the training on Day 2 focused primarily on practicing cyclic rounds, there were multiple types of rounds being called that differed from those seen in Day 1. While the nature of the rounds remains the same (e.g., they structured by litanies and similar crew behaviors are indicative of good performance), more data will be required to investigate this concept further.

When these aggregate graphs of volume are broken out by individual crewmember, similar patterns are seen. The graphs in Figure 11 (the more experienced team) display punctuated peaks in volume for the various crewmembers, particularly the Bearing Coordinator, ANAV, and Fathometer operator. The data displayed in Figure 12 contains more variance, which could have resulted from a combination of less strict adherence to the rhythm of the cyclic routines and more variation in the types of cyclic routines that were performed (Figure 12 does not display the QMOW due to an error in data collection that did not allow this graph to be made.) Again, while these differences suggest a means by which crew performance can be measured, more data is required to confirm this hypothesis.

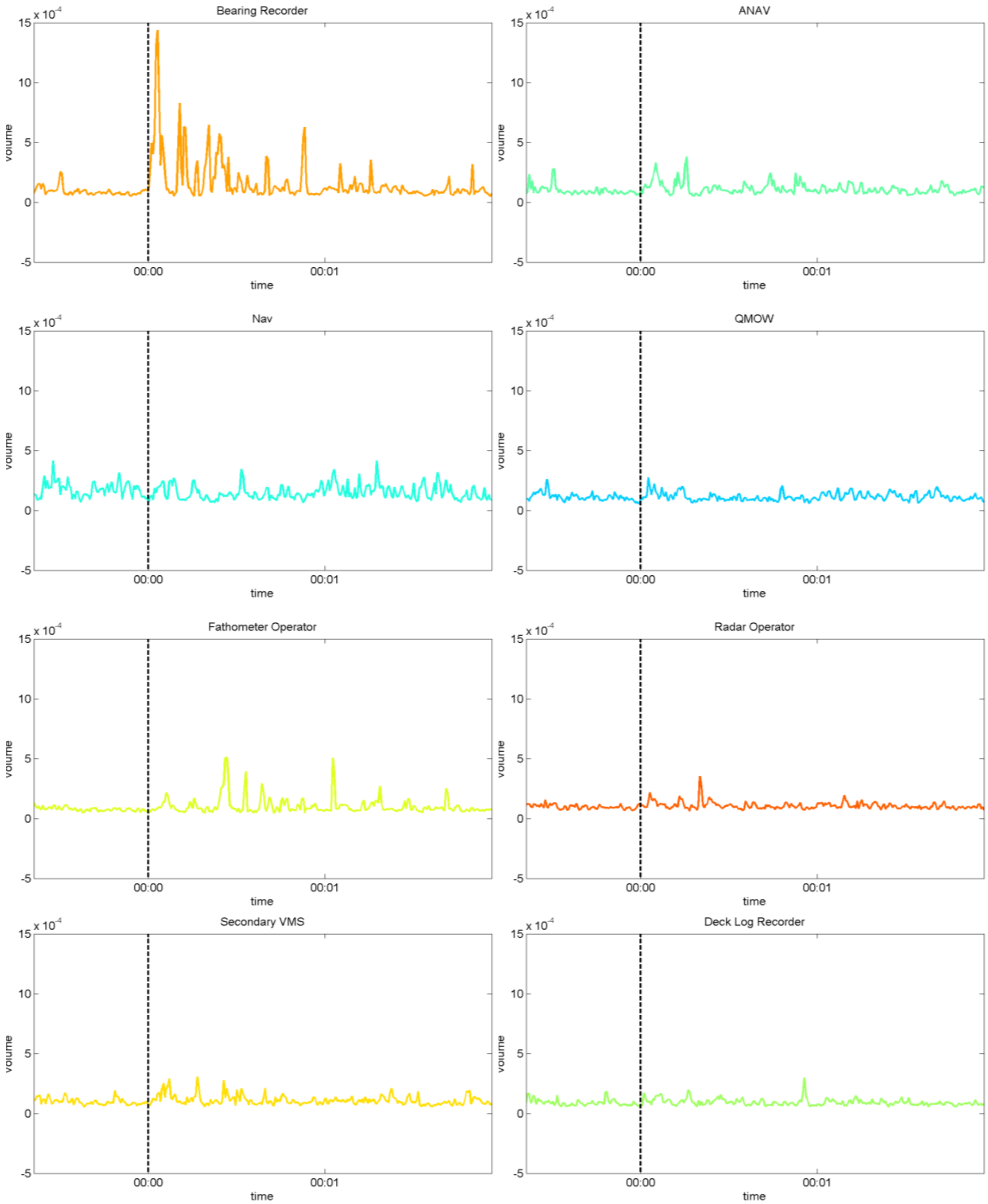

Figure 11: Average total volume for all cyclic routines by crewmember for Day 1 (Sessions 1 and 2). All volumes are measured in DBFS.

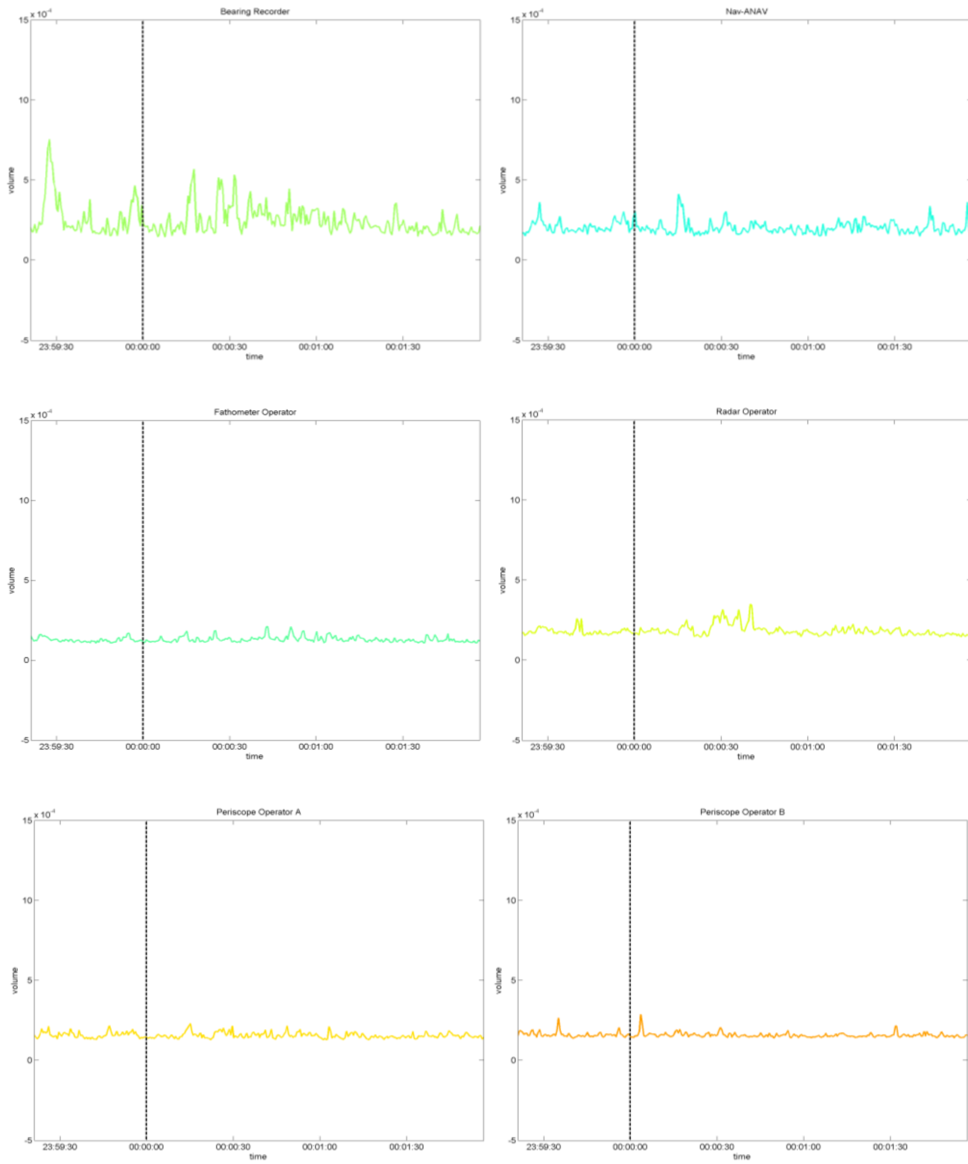

Figure 12: Average total volume for all cyclic routines by crewmember for Day 2 (Session 3). All volumes are measured in DBFS.

4.2 Infrared Data

The IR sensors on the badges detect when two badges are facing each other within a certain distance, a behavior which is assumed to correspond to a face-to-face interaction. The maximum distance at which a sensor can detect another is approximately 5-6 feet, and depending on how the badges are oriented with respect to one another, the distance of detection could be less. The IR sensors emit a beam every second, and when this beam is received by another badge, the identity of the emitting badge is recorded with a timestamp. Therefore, the total number of pings is an approximate measure of the amount of time that two badges were facing each other. Note that the angle at which an IR sensor can receive a signal is greater than the angle of the beam that is emitted, meaning that it is possible that one badge can record the presence of another without the other doing the same.

Figure 13 is a matrix that shows the frequency of interactions among the crew and workstations for Day 1 (Sessions 1 and 2). Each cell represents the number of interactions between the persons and/or equipment listed next to the corresponding row and column. “Hotter” colors correspond to more frequent interactions—note that the numbering on the legend to the right of matrix is an artifact of the scaling that was put in place to accentuate the differences between the cells. As a baseline check, none of the workstations interacted with one another, which can be seen in the block of blue cells to the upper left. The Nav and the ANAV each have numerous cells that show they interacted frequently with different crewmembers at various stations throughout the exercise. The station operators are seen interacting with their respective stations: e.g., the Fathometer Operator and the Fathometer, the Radar Operator and the Radar Station, and the Primary VMS Station and the Nav. Interestingly, the ANAV and the Secondary VMS Station interacted a number of times that was an order of magnitude greater than other interactions during the exercise. The ANAV was observed to spend much of his time at that station, but this relatively high number of pings could also have resulted from the manner in which the two badges were oriented.

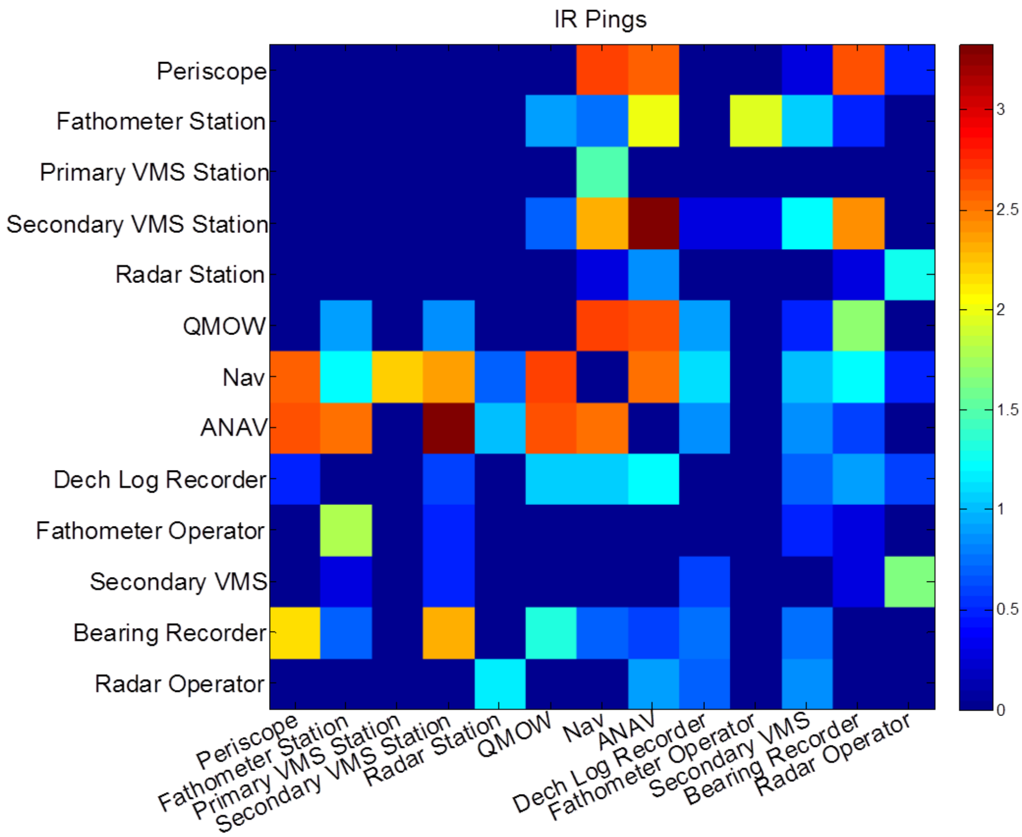

Figure 13: The frequency of interactions among crewmembers and workstations for Day 1, as recorded by the IR sensors. Hotter colors correspond to more frequent interactions.

Table 4 shows the raw number of pings for Day 1 for each interaction as seen in Figure 13 above. The cells that are blank did not have any registered pings. The matrix is not symmetrical because the design of the IR sensors allows a badge to receive a signal from another badge without the transmitting badge reciprocating the detection.

Table 4: The raw number of IR sensor pings recorded by each badge on Day 1. Blank cells did not have any registered pings.

	Periscope	Fathometer Station	Primary VMS Station	Secondary VMS Station	Radar Station	QMOW	Nav	ANAV	Dech Log Recorder	Fathometer Operator	Secondary VMS	Bearing Recorder	Radar Operator
Periscope							471	394			1	408	2
Fathometer Station						7	5	100		89	11	2	
Primary VMS Station							30						
Secondary VMS Station						4	213	2056	1	1	16	269	
Radar Station							1	6				1	18
QMOW		7		6			454	420	7		2	47	
Nav	368	15	159	243	4	498		339	12		9	15	2
ANAV	422	327		2132	9	418	328		6		6	3	
Dech Log Recorder	2			3		10	11	15			4	7	3
Fathometer Operator		58		2							2	1	
Secondary VMS		1		2					3			1	40
Bearing Recorder	148	4		208		20	4	3	5		5		
Radar Operator					13			7	4		6		

A similar interaction frequency matrix was generated for Day 2 (Figure 14). As in Figure 13, none of the stations registered interactions with each other, and the people who were assigned to different stations registered high numbers of interactions with their respective assignments. The Nav, who also served the role of the ANAV, registered the most varied interactions with the rest of the crew. In this case, the most interactions were recorded between the QMOW and the Bearing Recorder, again, with a number of interactions that was an order of magnitude higher than that of other crewmembers. They were observed to interact closely during the duration of the exercise, and in fact, they are required to coordinate closely as cyclic routines are performed. The second highest frequencies were seen with the Fathometer Operator’s interactions with the Fathometer Station and Primary VMS Station, which was expected given his position and the orientation of equipment in the trainer. Table 5 contains the raw data that is represented in Figure 14.

The patterns in each matrix do not suggest, at this time, any major differences between the teams that are more or less experienced. This could be in part because the trainers that were used on Day 1 and Day 2 were set up differently, and/or it could be due to different training objectives/focus. Also, the composition of each team was different, which may or may not affect one’s ability to detect differences in the graphs and attribute them to differences in team performance. However, this type of analysis may be useful in detecting deviations from expected behavior. For example, if the ANAV is observed to interact with certain crewmembers in a particular way during various missions, a change in this pattern could trigger a system to intervene (e.g., signaling to the ANAV that he might want to check a sensor that he has not checked in a while).

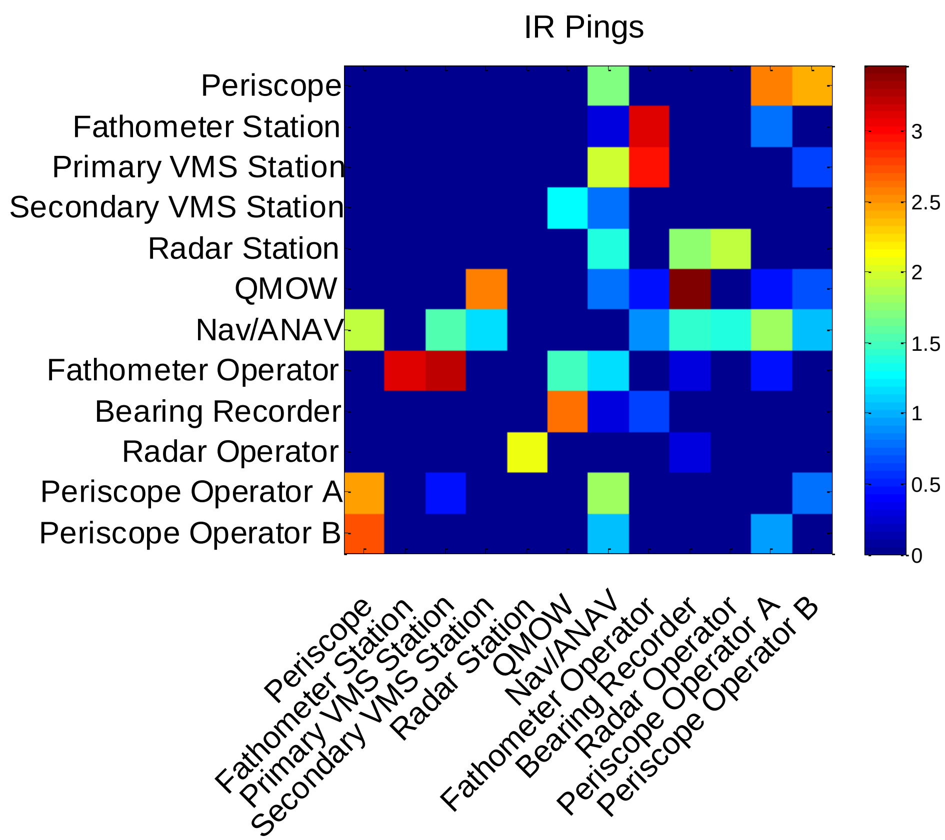

Figure 14: The frequency of interactions among crewmembers and workstations for Day 2, as recorded by the IR sensors. Hotter colors correspond to more frequent interactions.

Table 5: The raw number of IR sensor pings recorded by each badge on Day 2. Blank cells did not have any registered pings.

	Periscope	Fathometer Station	Primary VMS Station	Secondary VMS Station	Radar Station	QMOW	Nav/ANAV	Fathometer Operator	Bearing Recorder	Radar Operator	Periscope Operator A	Periscope Operator B
Periscope							49				366	254
Fathometer Station							1	1334			5	
Primary VMS Station							89	882				3
Secondary VMS Station						17	5					
Radar Station							23		57	83		
QMOW				384			5	2	2925		2	4
Nav/ANAV	81		33	13				7	26	23	61	11
Fathometer Operator		1371	1617			31	14		1		2	
Bearing Recorder						397	1	3				
Radar Operator					120				1			
Periscope Operator A	279		2				66					5
Periscope Operator B	569						10				8	

The IR sensor data was also used to show where each crewmember tended to spend his time in the control room (Figures 15 and 16). While a different representation of the same data, it does show more clearly how each person distributed his time across different locations. Those who were assigned to different stations tended to spend most of their time there, while the Nav and ANAV were typically more interactive and spent more time at 2-3 different stations. Once again, it is difficult to make firm conclusions about differences that can be seen between the two crews, but more investigation may show how these data tend to change over time given a particular crew and control room configuration.

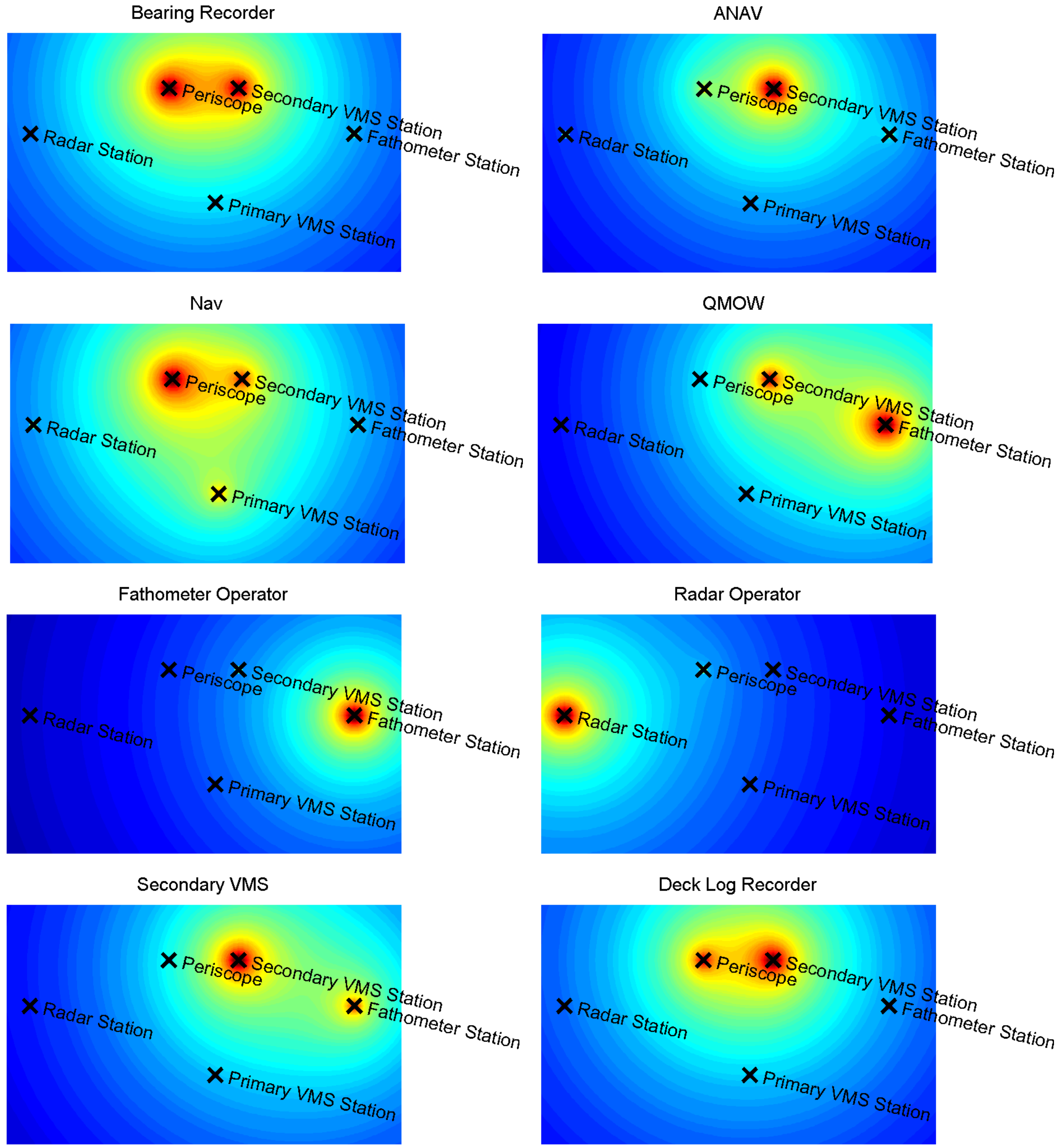

Figure 15: The IR sensor data was used to plot recorded position over time of all crewmembers within the control room on Day 1 (Sessions 1 and 2). Hotter colors correspond to more time spent in a particular area.

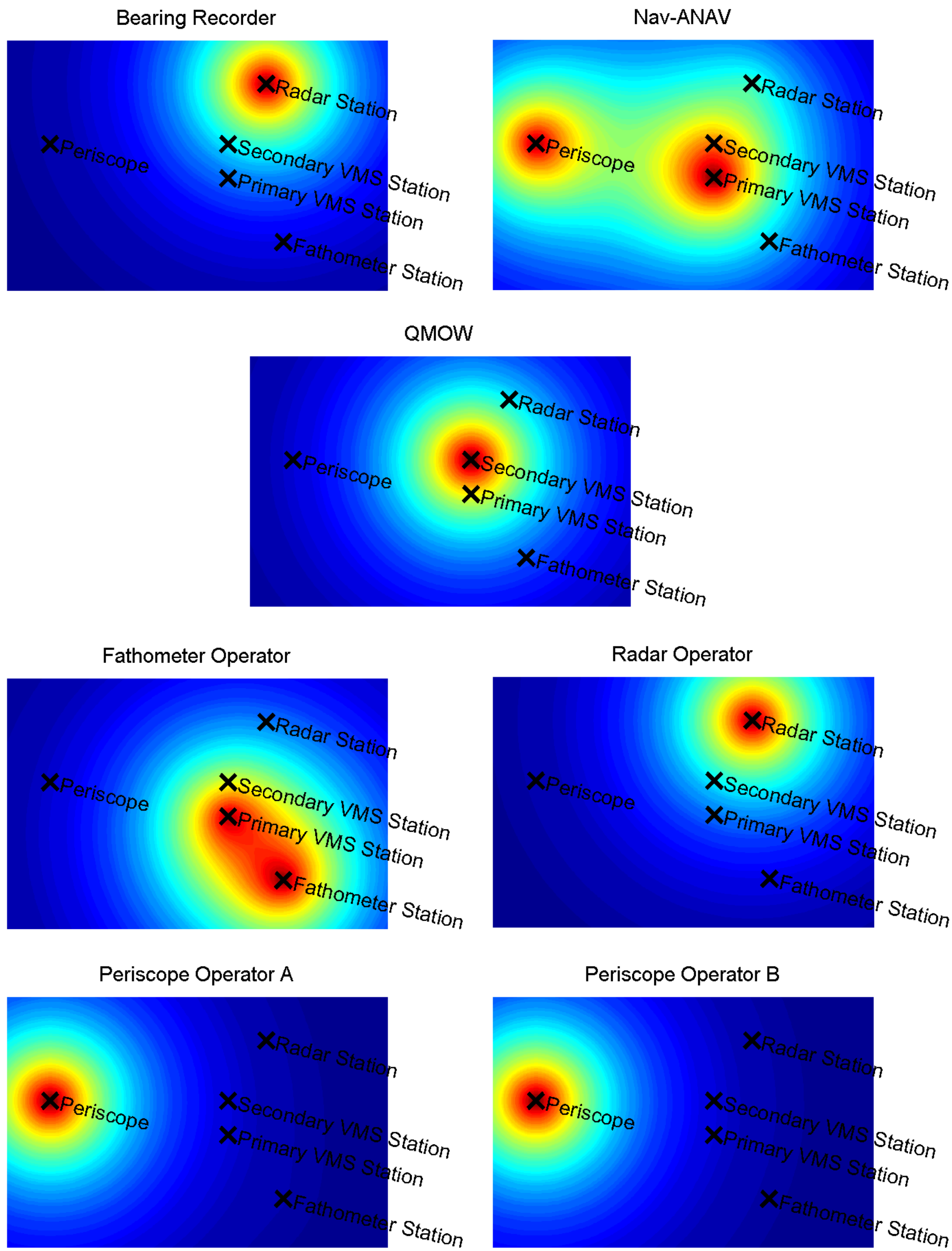

Figure 16: The IR sensor data was used to plot recorded position over time of all crewmembers within the control room on Day 2 (Session 3). Hotter colors correspond to more time spent in a particular area.

4.3 Bluetooth Data

As discussed earlier, the Bluetooth transceiver on each badge sends a signal on average every 5 minutes, and then records the identities of the badges that received this signal and sends back a reply. In addition to being able to tell when badges were close enough to receive and send this signal, the Received Signal Strength Indicator (RSSI) value is also recorded. This value corresponds to the strength of the connection between the two badges, which is hypothesized to correspond to the distance between them. Initial tests at SNL concluded that this particular feature may not be sensitive enough for employment within the setting studied here, and through further exploration at NSS, our data show those that initial findings were generally confirmed.

Given that every badge sends out Bluetooth “pings” and records the identity of badges that respond, if the RSSI values recorded by two badges was a reliable indicator of the distance between them, then one would expect that these recorded values would be correlated. That is, as the RSSI value with respect to badge “B” increases as recorded by badge “A,” then the RSSI value with respect to badge “A” as recorded by badge “B” should also increase To investigate this empirically, signal strength (which was recorded at roughly 5-minute intervals) was calculated over time by linearly interpolating between the recorded values. If at any given point in time the signal strengths were identical, then there would be a strong linear correlation between the two. Figure 17 shows the best linear fit for the Nav and the Periscope Station for Day 1. There appears to be no correlation as seen in the scatter plot. This was confirmed by graphing the cross correlation between the two strength signals over time (the graph below the scatter plot). Unfortunately, within the range of distances studied here, these data suggest that the data are too noisy to be reliable indicators of distance between the Nav and the Periscope Station.

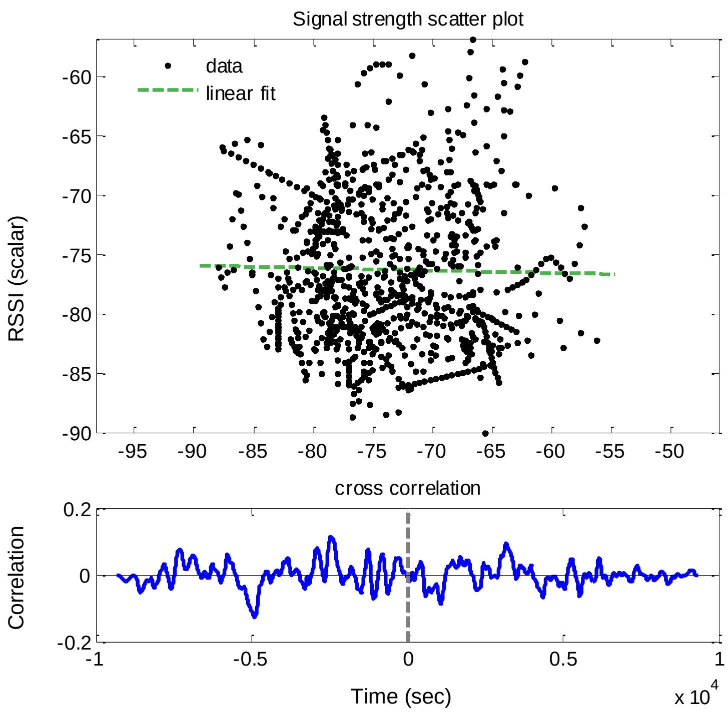

Figure 17: Bluetooth data correlation plot between the Nav and the Periscope Station for Day 1.

Although the Nav served as the Periscope Operator during the exercise, the fact that he tended to move around the control room could have contributed to the noise in the data. However, similarly low correlations are seen between other operators and their assigned workstations (e.g., Figures 18 and 19).

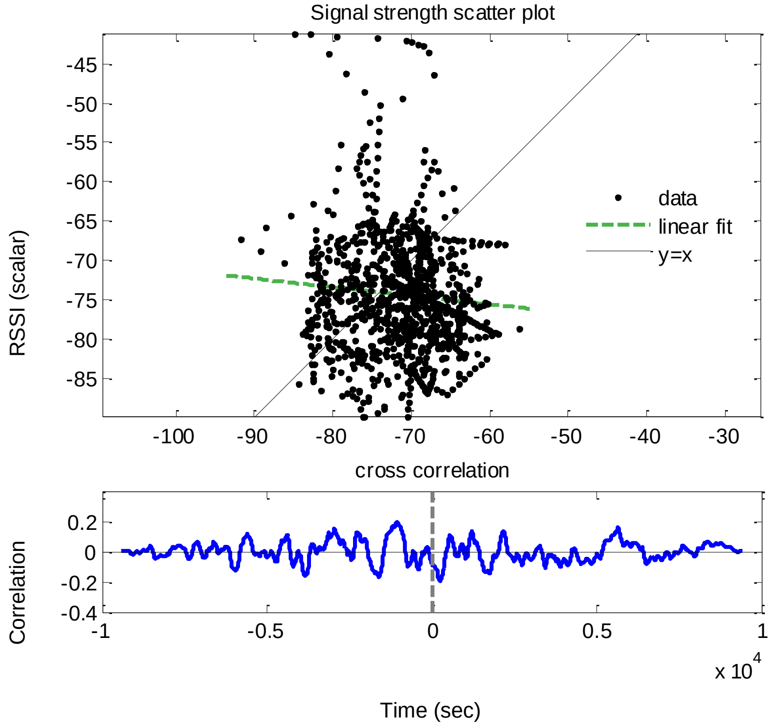

Figure 18: Bluetooth data correlation plot between the Fathometer Operator and Fathometer Station - Day 1.

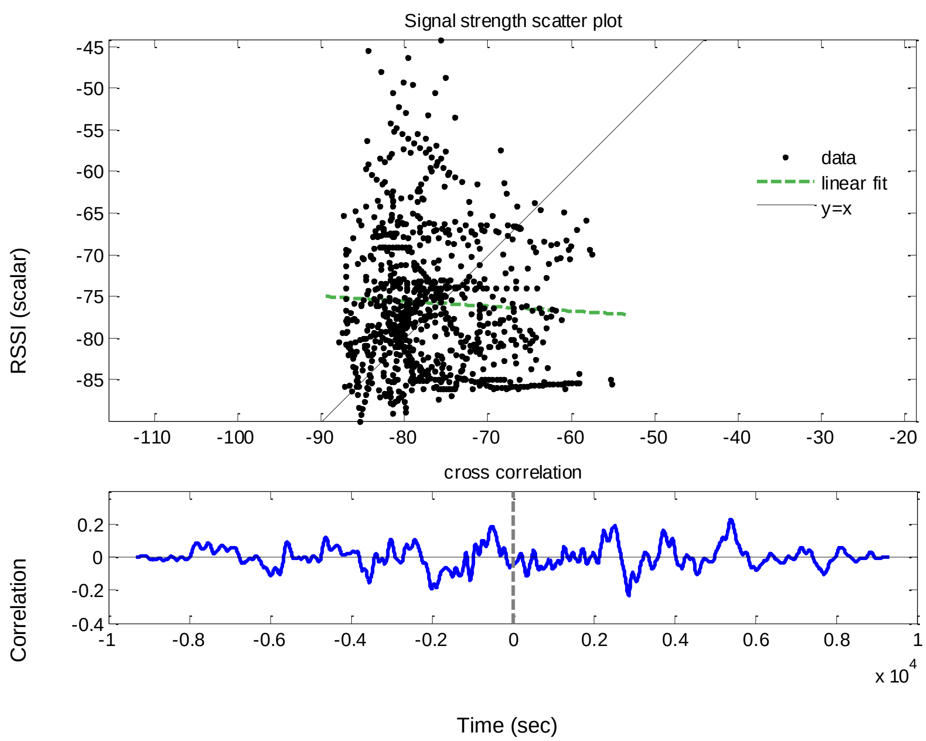

Figure 19: Bluetooth data correlation plot between the Radar Operator and Radar Station - Day 1.

However, when the correlations are explored over a longer length of time, there are some differences that can be seen. Figure 20 shows the Bluetooth correlation data between the Radar Operator and the Radar Station for the entire first day (including time before, during, and after the exercise, for as long as both badges were turned on). It is unclear what happened beyond the training session, but it is likely that the badges were separated by a distance much greater than would have been experienced during the scenario (e.g., by taking the two badges into separate rooms). This suggests that the sensitivity of the Bluetooth signal alone may not currently be adequate to detect the relative positions of crewmembers in this environment. However, exercises in other domains may find these current capabilities sufficient if they take place across longer distances. Furthermore, as the technology becomes refined, or as new technology is used to measure signal strength and infer separation, more reliably correlated data will enable new analyses.

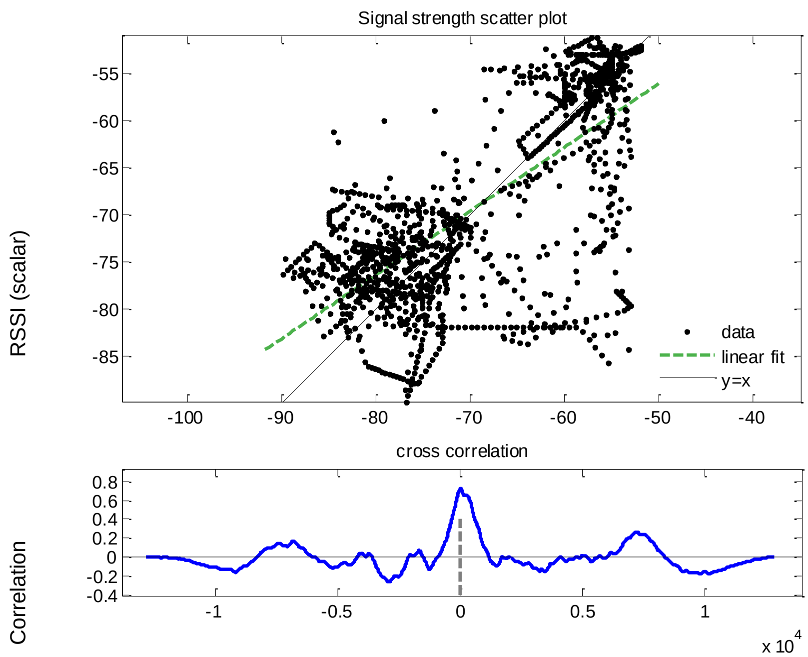

Figure 20: Bluetooth data correlation plot between the Radar Operator and Radar Station over the course of the entire Day 1 (data was recorded before, during, and after training).

4.4 Self-Reported Performance

As the crewmembers returned their badges at the conclusion of each exercise, each was asked to complete a survey that asked several questions intended to capture how they felt they performed individually and as a team (see *Appendix B: Self-Reported Performance Survey*). The questions rated on a 5-point Likert scale with half-point increments, with “1” corresponding to “poor” performance and “5” corresponding to “outstanding” performance. The five questions that were asked are:

1. How unified do you feel the team performed during the exercise?
2. Overall, how would you rate the performance of the entire team?
3. Overall, how you would rate your performance during the exercise?
4. Overall, how well did the team do in minimizing cross-track error?
5. Overall, how well did the team do with respect to maintaining safety?

For Day 1, there were eight crewmembers who completed the survey: the Bearing Recorder, Radar Operator, QMOW, ANAV, Nav, Fathometer Operator, Deck Log Recorder, and Secondary VMS Operator. For Day 2, the survey was completed by the Bearing Recorder, Radar Operator, QMOW, ANAV, Fathometer Operator, Periscope Operator A, and Periscope Operator B. Table 6 displays the average response for each question for Days 1 and 2.

Table 6. Self-Reported performance by question for Days 1 and 2.

Question:	Day 1	Day 2
1. How unified do you feel the performed during the exercise?	4.2	3.1
2. Overall, how would you rate the performance of the entire team?	4.2	2.7
3. Overall, how you would rate your performance during the exercise?	3.7	3.2
4. Overall, how well did the team do in minimizing cross-track error?	3.9	3.1
5. Overall, how well did the team do with respect to maintaining safety?	4.5	3.8

On average, the crew who performed during Day 2 did not feel as if they performed as well as the crew on Day 1 felt they did. While both teams were intact (i.e., they performed together as a Watch Section on their respective ships), the crew on Day 1 was more experienced than the crew on Day 2. The lack of experience could explain the lower ratings across every question. This difference can be seen more clearly in Figure 21. The greatest spread in average response occurred with Question 2, which focused on how each individual would rate the performance of the entire team. The experienced team (Day 1) rated themselves quite high, 4.2 on average, while the less experienced team rated themselves at 2.7 on average. This was also the lowest rating recorded for Day 2. The highest rating for each crew was associated with Question 5, which asked how well each individual thought the team performed with respect to maintaining the safety of the ship. In conclusion, the self-reported survey captures differences between the performances of each team, which allows us to look for other correlations in the badge data. For example, given that the team on Day 1 felt as if they were more unified than was reported by the team on Day 2, future analysis Sociometric Badge data may want to focus on indicators of that correlate to the team's sense of unity. Supplementing data collection with additional sources of information enhances the range of conclusions that can be derived, and therefore improves our interpretation of assessment of crew performance.

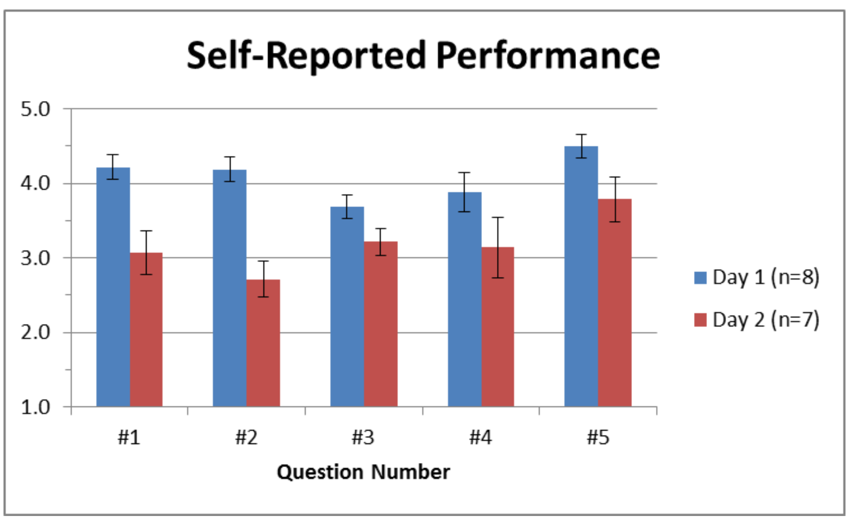

Figure 21: Self-Reported performance by question for Days 1 and 2 (the error bars indicate standard error).

5. Conclusion and Future Work

In conclusion, the Sociometric Badges continue to be a promising solution to augment the data collection methods being employed in the AT-SNAP program. They are an unobtrusive, passive means of automatically collecting a variety of data that can be analyzed to assess team processes in the absence of instructors or observers. By using a unique combination of sensors, team behavior can be examined among multiple dimensions that have not been examined to the extent that they now can be. This exploratory effort suggests that the badges are likely able to diagnose aspects of team performance during navigation exercises, and numerous preliminary findings suggest that the badge technology is well-suited to the undersea warfare domain.

The volume level in the control room appears to be an indicator of tension, and patterns in individual volume seem to correlate to interesting behaviors. For example, it is possible to identify when commands are given and the order in which two crewmembers spoke. In future research efforts, it may be possible to use this data to identify more nuanced behaviors, such as individual leadership styles. Departures from an observed baseline volume are expected at certain times, such as when the crew is faced with warnings from the system (e.g., a yellow sounding). The volume data can be used to determine whether or not the crew is acting in an expected or appropriate way. The data is also useful when examining the manner in which cyclic routines are performed. There are differences in the patterns of data that were collected from the two teams that seem as if they could be used to determine the experience level of each. Overall, these data suggest that volume as captured by the Sociometric Badges may be a promising way to detect what the team is doing (e.g., where their focus of attention is) and determine what they should be doing (e.g., patterns in volume that correspond to better execution of cyclic routines; tension that should exist given certain mission conditions).

The IR sensor data provide a rough picture of how crewmembers interact both with one another and various workstations. The number of different people with whom crewmembers interact, and the frequency of those interactions, can be easily captured and graphed for analysis. Preliminary results suggest that the IR data can be used to map of control room activity which can be used to compare behavior of more and less experienced teams. However, patterns in IR sensor data may be specific to the control room configuration and to the crew configuration. In future work, the IR sensor data may be more useful in determining changes in behavior within a crew in a consistent environment, rather than between two entirely different teams and settings. These data can also be used to plot graphs that show where a crewmember tended to spend most of his time within the control room. This representation can be used to visualize how a crewmember moves within the space, and with more data, could be found to correlate with performance.

Some challenges remain, but as the technology matures, there will be additional opportunities to advance these diagnostic capabilities even further. For example, the Bluetooth signal is currently not as sensitive as it needs to be in this environment to be able to reliably determine the distance between crewmembers based on the RSSI value. However, as this technology evolves or as different/new technology is used, this accuracy will increase. In addition, there are additional features being developed, and new analyses that are being refined, that will further explore the benefits of the Sociometric Badges in future efforts. For example, the energy data that was collected is unexplored, the volume data can be further analyzed to derive quantitative measures that characterize conversation, and additional interpretations of the data can be applied to reduce noise and identify relevant patterns and additional indicators of team skill. To summarize, the

Sociometric Badges are novel, promising, and an exciting next step in automated submarine team assessment.

6. References

1. Chester, S., Jones, E., Jackson, C., & Diedrich, F. (2011). *Future Naval Capabilities Measures Support: Evolution of Measure Hierarchy*. Technical Report, Aptima, Inc., Woburn, MA.
2. Jones, E., Diedrich, F., Jackson, C., Armbruster, R., & Steed, R. (2010). *Exceptional Expertise for Submarine Command Team Decision Making (E2SCDM)*. Technical Report, Aptima, Inc., Woburn, MA.
3. Olguin Olguin, Daniel, and Alex Pentland. "Assessing Group Performance from Collective Behavior." CSCW 2010 Workshop on Collective Intelligence in Organizations. Savannah, GA USA, 2010.
4. Smith-Jentsch, K.A., Zeisig, R.L., Acton, B., & McPherson, J.A. (1998). Team dimensional training: A strategy for guided team self-correction. In Cannon-Bowers & Salas (Eds.), *Making decisions under stress: Implications for individual and team training*. Washington, DC: American Psychological Association.
5. Sociometric Solutions, Incorporated (2011). *Sociometric Badge 03-02, Preliminary User Guide, Revision 1.6*.

1. LIST OF ACRONYMS, ABBREVIATIONS, AND SYMBOLS

ACRONYM	DESCRIPTION
ANAV	Assistant Navigator
AT-SNAP	Adaptive Training for Submarine Navigation and Piloting
CTSS	Continuing Training Support System
DBFS	Decibels Relative to Full Scale
GPS	Global Positioning Satellite System
IR	infrared
Nav	Navigator
NSS	Naval Submarine School
ONR	Office of Naval Research
QMOW	Quarter Master of the Watch
ROC	round of contacts
RSSI	Received Signal Strength Indication
SNL	Sandia National Laboratories
SPAN	Submarine Piloting and Navigation
SSI	Sociometric Solutions, Inc.
TDT	Team Dimensional Training
VMS	Voyage Management System
XTE	cross-track error

Appendix A: Experimentation Forms

ATSNAP Badge Assignment Worksheet

Date: ____________________ *Time Start:* ______________________

Session: __________________ *Time End:* _______________________

QMOW	
Nav	
ANAV	
Radar	
Fathometer Operator	
Tech Log Recorder	
Other:____________	
Periscope	
VMS Station	
Radar Station	
Fathometer Station	

Cross-Track Error

Date: ______________________ *Time Start:* _______________________

Session: ____________________ *Time End:* ________________________

Time	Cross-Track Error
0:05	
0:10	
0:15	
0:20	
0:25	
0:30	
0:35	
0:40	
0:45	
0:50	
0:55	
1:00	
1:05	
1:10	
1:15	
1:20	
1:25	
1:30	
1:35	
1:40	
1:45	
1:50	
1:55	
2:00	
2:05	

2:10	
2:15	
2:20	
2:25	
2:30	
2:35	
2:40	
2:45	
2:50	
2:55	
3:00	
3:05	
3:10	
3:15	
3:20	
3:25	
3:30	
3:35	
3:40	
3:45	
3:50	
3:55	
4:00	

Course Changes

Date: ______________________ *Time Start:* _______________________

Session: ____________________ *Time End:* ________________________

Time	**Course Change**	**Notes**

Round of Contacts

Date: ______________________ *Time Start:* _______________________

Session: ____________________ *Time End:* ________________________

#	Time Start	Time End
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		

Scenario Events

Date: ______________________   *Time Start: _______________________*

Session: ____________________   *Time End: ________________________*

Time	Events	SME Comments

Appendix B: Self-Reported Performance Survey

Self-Reported Performance

How unified do you feel the team performed during the exercise?

1	2	3	4	5
Not at all cohesive		Neutral		Very cohesive

Overall, how would you rate the performance of the **entire** team?

1	2	3	4	5
Poor		Average		Outstanding

Overall, how would you rate **your** performance during the exercise?

1	2	3	4	5
Poor		Average		Outstanding

Overall, how well did the team do in minimizing cross-track error?

1	2	3	4	5
Poor		Average		Outstanding

Overall, how well did the team do with respect to maintaining safety?

1	2	3	4	5
Poor		Average		Outstanding

Clearance Request for Public Release of Department of Defense Information	See instructions on reverse	ONRINST 5570.2A

To: Chief of Naval Research (Code BD043)	CASE NO. 43-362-12

I. Document Description *(To be completed by author)*

1. TYPE OF DOCUMENT	2. TITLE
Final Report	Performance Measures for Improved Submarine Decision Making
3. PAGE COUNT	4. SUBJECT AREA
47	ONR BAA11-001

II. Author/Speaker *(To be completed by author)*

5. NAME *(Author/Speaker)*	6. RANK/GRADE	7. OFFICE/CODE
Eric Jones, Jonathan Lansey, F	N/A	N/A

8. TITLE *(Author/Speaker)*	9. ACTIVITY
N/A	Upload Final Report

III. Presentation/Publication Data *(To be completed by author)*

10. OCCASION *(Title of Meeting/Publication for which Intended)*	11. MEANS *(Means by which material will be presented)*
Defense Technical Information Center (DTIC) Website; Any oth	Available on DTIC Website; TBD
12. DATE *(Of meeting or publication date and issue)*	13. PLACE *(Location of meeting or where published)*
15 Jun 2012	Virtual; TBD
14. SPONSOR *(Organization conducting meeting or publishing material)*	15. CONFERENCE TYPE *(Check R of following)*
Defense Technical Information Center	☐ CLOSED FORUM ☒ OPEN FORM
16. IF MATERIAL IS PRESENTED, WILL IT BE PRINTED?	17. IF PRINTED, SPECIFY PUBLICATION NAME
☒ YES ☐ NO	TBD

IV. Control Data *(To be completed by author) (Check "yes" or "no")*

18. INFORMATION CONTAINED IN PRESENTATION OR PUBLICATION IS CONSIDERED:
(Indicate answer by placing an "X" in proper column.)

	YES	NO
a. REGARDED AS UNCLASSIFIED	X	
b. SUBJECT TO INTERNATIONAL ARMS REGULATIONS (ITAR)		X
c. CONTAINS MILITARILY CRITICAL TECHNOLOGY		X
d. PROPRIETARY DATA		X
e. RESTRICTED BY PATENT		X
f. REQUIRES OTHER APPROVAL. IF YES, STATE BY WHOM ______		X
g. PREPARED UNDER ONR CONTRACT? IF YES, STATE CONTRACT NO N00014-11-M-0329	X	
h. CONTRACT IS CLASSIFIED		X
i. ALL REFERENCES ARE UNCLASSIFIED	X	
(1) IF NO, STATE REFERENCES THAT ARE CLASSIFIED ______		

(2) STATE RELEASE AUTHORITY FOR CLASSIFIED REFERENCES ______

19. PROGRAM MANAGER/SPONSOR NAME:	20. CODE
Dr. William Krebs	342

NAVONR 5510/7 (Rev. 09-07)

Security
JUN 0 5 2012

RECEIVED
JUN 0 8 2012

V. Coordination/Approvals

21. *(Indicate Yes or No and show Initials, Code and Date)*
(Note: "a" only for contractor papers, "b" through "e" for ONR employee papers)

	YES	NO	CODE	INITIALS	DATE
a. CONTRACT MONITOR		X			
b. PROGRAM MANAGER	X		342	[illegible]	5 JUN 2012
c. DIVISION DIRECTOR		X			
d. DEPARTMENT/DIRECTORATE DIRECTOR		X			
e. PUBLIC AFFAIRS	X		CSC	[illegible]	6/8/12

(If no, give reason for:)

22. RECOMMEND DISTRIBUTION STATEMENT *(check one)*

A B C D E X

VI Clearance Action *(To be completed by Staff Security Branch (Code 00R1)*

23. SIGNATURE AND TITLE OF STAFF SECURITY BRANCH REPRESENTATIVE APPROVING FOR ONR

[illegible signature]

24. DATE

7 Jun 12

23a. Staff Security Branch Represenative check approved or disapproved:

☑ APPROVED ☐ DISAPPROVED

25. DATE SUBMITTED TO CNO (OP-09N)

26. LTR SERIAL NO.

27. CNO (OP-09N)

☐ APPROVED ☐ DISAPPROVED

28. DATE

29. SERIAL NO.

30. REASON FOR DISAPPROVAL BY SNO (OP-09N)

31. DATE AUTHOR/PRESENTER NOTIFIED

32. SERIAL NO OF NOTIFICATION LTR

NAVONR 5510/7 (Rev. 09-07)